



PTO/SB/08A (10-01)

Approved for use through 10/31/2002. OMB 0651-0031

U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449A/PTO				<i><b>Complete if Known</b></i>	
<b>INFORMATION DISCLOSURE STATEMENT BY APPLICANT</b>				Application Number	10/676,918
(use as many sheets as necessary)				Filing Date	09-30-2003
				First Named Inventor	NADAV EIRON
				Art Unit	
				Examiner Name	NOT ASSIGNED
				Attorney Docket Number	ARC920030028US1
Sheet	1	of	4		

#### **U.S. PATENT DOCUMENTS**

## **FOREIGN PATENT DOCUMENTS**

Examiner Signature		Date Considered	
-----------------------	--	--------------------	--

**\*EXAMINER:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

<sup>1</sup> Applicant's unique citation designation number (optional). <sup>2</sup> See Kinds Codes of USPTO Patent Documents at [www.uspto.gov](http://www.uspto.gov) or MPEP 901.04. <sup>3</sup> Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). <sup>4</sup> For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. <sup>5</sup> Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST. 16 if possible. <sup>6</sup> Applicant is to place a check mark here if English language Translation is attached.

Burden Hour Statement: This form is estimated to take 2.0 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Alexandria VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, Alexandria VA 22313-1450.



PTO/SB/08B (10-01)

Approved for use through 10/31/2002. OMB 0651-0031  
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449B/PTO

## INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

Sheet

4 of 4

### Complete if Known

Application Number	10/676,918
Filing Date	09-30-2003
First Named Inventor	NADAV EIRON
Group Art Unit	
Examiner Name	NOT ASSIGNED
Attorney Docket Number	ARC920030028US1

### OTHER PRIOR ART -- NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No. 1	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published	T <sup>2</sup>
		[1] HTML 4.01 specification, W3C recommendation, December 1999.	
		[2] Lada A. Adamic and Eytan Adar. Friends and neighbors on the web. <a href="http://www.hpl.hp.com/shl/people/eytan/fnn.pdf">http://www.hpl.hp.com/shl/people/eytan/fnn.pdf</a> .	
		[3] Tim Berners Lee, James Hendler, and Ora Lassila. The semantic web. Scientific American, May 2001, <a href="http://www.sciam.com/2001/0501issue/0501berners-lee:html">http://www.sciam.com/2001/0501issue/0501berners-lee:html</a>	
		[4] Rodrigo A. Botafogo and Ben Shneiderman. Identifying aggregates in hypertext structures. In UK Conference on Hypertext, pages 63-74, 1991.	
		[5] Vannevar Bush. As we may think. The Atlantic Monthly, 176(1):101-108, July 1945.	
		[6] Michael Chen, Marti A. Hearst, Jason Hong, and James Lin.. Cha-cha: A system for organizing intranet search results. In USENIX Symposium on Internet Technologies and Systems, 1999.	
		[7] R. C. Daley and P. G. Neumann. A general purpose file system for secondary storage. In AFIPS Conference Proceedings, volume 27, pages 213-229, 1965.	
		[8] M. M. Kessler. Bibliographic coupling between scientific papers. American Documentation, 14, 1963.	
		[9] S. Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, and Andrew Tomkins. Extracting large scale knowledge bases from the Web. In Proceedings-of-the-25th-VLDB-Conference, pages 639-650, 1999.	
		[10] Yoshiaki Mizuuchi and Keishi Tajima. Finding context paths for Web pages. In Proceedings of Hypertext 99, pages 13-22, Darmstadt, Germany, 1999.	

Examiner Signature	Date Considered
--------------------	-----------------

\*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

<sup>1</sup> Applicant's unique citation designation number (optional). <sup>2</sup> Applicant is to place a check mark here if English language Translation is attached.

Burden Hour Statement: This form is estimated to take 2.0 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Alexandria VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, Alexandria VA 22313-1450.



PTO/SB/08B (10-01)

Approved for use through 10/31/2002. OMB 0651-0031

U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449B/PTO

## INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

Sheet 3 of 4

**Complete if Known**

Application Number	10/676,918
Filing Date	09-30-2003
First Named Inventor	NADAV EIRON
Group Art Unit	
Examiner Name	
Attorney Docket Number	ARC920030028US1

**OTHER PRIOR ART -- NON PATENT LITERATURE DOCUMENTS**

Examiner Initials'	Cite No. 1	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published	T <sup>2</sup>
		[11] Theodor Holm Nelson. Embedded markup considered harmful. <a href="http://www.xml.com/pub/a/w3j/s3.nelson.html">http://www.xml.com/pub/a/w3j/s3.nelson.html</a> , October 1997.	
		[12] David M. Pennock, Gary W. Flake, Steve Lawrence, Eric J. Glover, and C. Les Giles. Winners don't take all: Characterizing the competition for links on the web. Proceedings of the National Academy of Science, 99(8):5207-5211, April 16 2002.	
		[13] Craig Silverstein, Monika Henzinger, Hannes Marais, and Michael Moricz. Analysis of a very large altavista query log. Technical report, DEC Systems Research Center, 1998, Technical note 1998 14.	
		[14] H. A. Simon. On a class of skew distribution functions. Biometrika, 42(3/4):425-440, 1955.	
		[15] H. G. Small. Co citation in the scientific literature: A new measure of the relationship between two documents..Journal-of-American-Society-for-Information-Science,-24(4):265-269,-1973.	
		[16] Ellen Spertus. Parasite: Mining structural information on the web. In Proceedings of the Sixth Internation Conference on the World Wide Web, volume 29 of Computer Networks, pages 1205-1215, 1997.	
		[17] Amanda Spink, Dietmar Wolfram, B. J. Jansen, and Tefko Saracevic. Searching the web: The public and their queries. Journal of the American Society for Information Science, 53(2):226-234, 2001.	
		[18] Keishi Tajima, Kenji Hatano, Takeshi Matsukura, Ryouichi Sano, and Katsumi Tanaka. Discovery and retrieval of logical information units in web. In Proceedings of the Workshop on Organizing Wep Space (WOWS 99), pages 13-23, Berkeley, CA, August 1999.	
		[19] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of "small world networks". Nature, 393:440-442, June 4 1998.	
		[20] Ron Weiss, Bienvenido Velez, Mark A. Sheldon, Chanathip Namprempre, Peter Szilagyi, Andrzej Duda, and David K. Gifford. Hypersuit: A hierarchical network search engine that exploits content link hypertext clustering. In ACM Conference on Hypertext, pages 180-193, Washington USA, 1996.	
		[21] Nadav Eiron, Kevin McCurley. Locality, Hierarchy, and Bidirectionality in the Web, <a href="http://www.almaden.ibm.com/cs/people/mccurley/pdfs/graph.pdf">http://www.almaden.ibm.com/cs/people/mccurley/pdfs/graph.pdf</a>	
		[22] Gary Flake, Steve Lawrence, C. Lee Giles. Efficient Identification of Web Communities, Proceedings of the Sixth International Conference on Knowledge Discovery and Data Mining (ACM SIGKDD-2000), p. 150-160, August 20-23, 2000, Boston, MA, USA.	

Examiner Signature	Date Considered
--------------------	-----------------

\*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

<sup>1</sup> Applicant's unique citation designation number (optional). <sup>2</sup> Applicant is to place a check mark here if English language Translation is attached.

Burden Hour Statement: This form is estimated to take 2.0 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Alexandria VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, Alexandria VA 22313-1450.



PTO/SB/08B (10-01)

Approved for use through 10/31/2002. OMB 0651-0031  
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449B/PTO

## INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

Sheet

4 of 4

### Complete if Known

Application Number	10/676,918
Filing Date	09-30-2003
First Named Inventor	NADAV EIRON
Group Art Unit	2131
Examiner Name	NOT ASSIGNED
Attorney Docket Number	ARC920030028US1

### OTHER PRIOR ART -- NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No. 1	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T <sup>2</sup>
		[23] Keishi Tajima, Yoshiaki Mizuuchi, Masatsugu Kitagawa, Katsumi Tanaka. Cut-as-a-Querying-Unit-for-WWW-Netnews-and-E-Mail-Proc-Of-ACM-Hypertext'98, June 1998, Pittsburgh, PA, USA.	
		[24] Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, Andrew Tomkins. Trawling the web for emerging cyber-communities, Computer Networks, v. 31, n. 11-16, p. 1481-1493, 1999, Amsterdam, Netherlands.	
		[25] Xiaofeng He, Chris H.Q. Ding, Hongyuan Zha, Horst D. Simon. Automatic Topic Identification Using Webpage Clustering, ICDM 2001, Proceedings IEEE Conference on Data Mining, 2001, p. 195-202.	
		[26] Donna Bergmark. Models and tools for generating digital libraries: Collection synthesis, Proceedings of the second ACM/IEEE-CS joint conference on Digital libraries, July 2002, p. 253-262.	
		[27] Activity analysis of real world entities by combining dynamic information sources and real world entities, Research Disclosure, n. 433100, May 2000, p. 905-906.	
		[28] Virtual URLs for Browsing and Searching Large Information Spaces, Research Disclosure, n. 41392, September 1998, p. 1238-1239.	
		[29] Database for Navigational Links for a Text Part in a Compound Document, IBM Technical Disclosure Bulletin, v..38, n..1, January 1995, p..371-372.	
		[30] Lada A. Adamic. The Small World Web, Proc. 3rd European Conf. Research and Advanced Technology for Digital Libraries (ECDL), n. 1696, 1999, p. 443-452, Springer-Verlag, S. Abiteboul and A.-M Vercoustre eds. a third code means for processing-the-compound-document,-wherein-the-processing-includes-at-least-one of creating at least one presentation of results from retrieval, summarization, and classification.	

Examiner Signature	Date Considered
--------------------	-----------------

\*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

<sup>1</sup> Applicant's unique citation designation number (optional). <sup>2</sup> Applicant is to place a check mark here if English language Translation is attached.

Burden Hour Statement: This form is estimated to take 2.0 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Alexandria VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Alexandria VA 22313-1450.